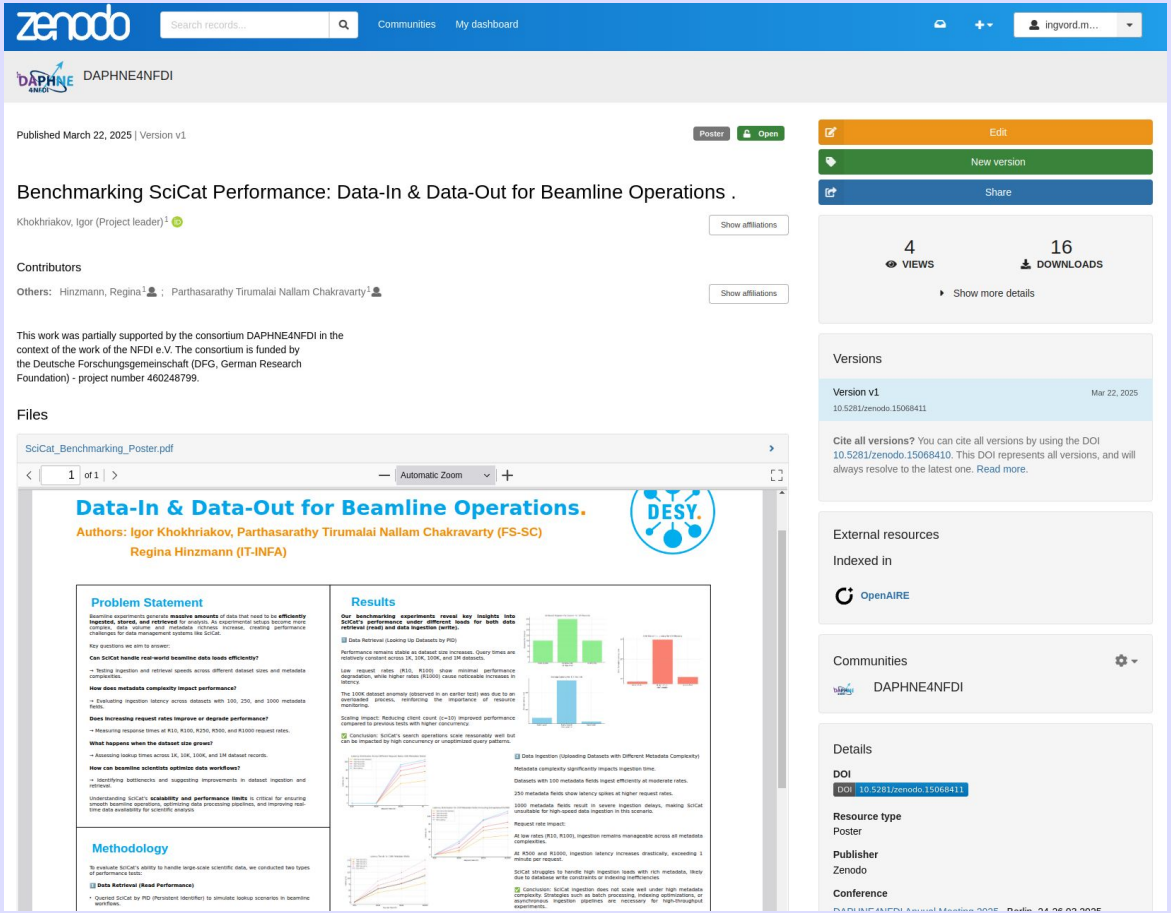
SciCat Performance Assessment

SciCatCon 2025

Igor Khokhriakov aka Ingvord

MORE DETAILS IN THE DAPHNE4NFDI ANNUAL MEETING POSTER

https://zenodo.org/records/15068411



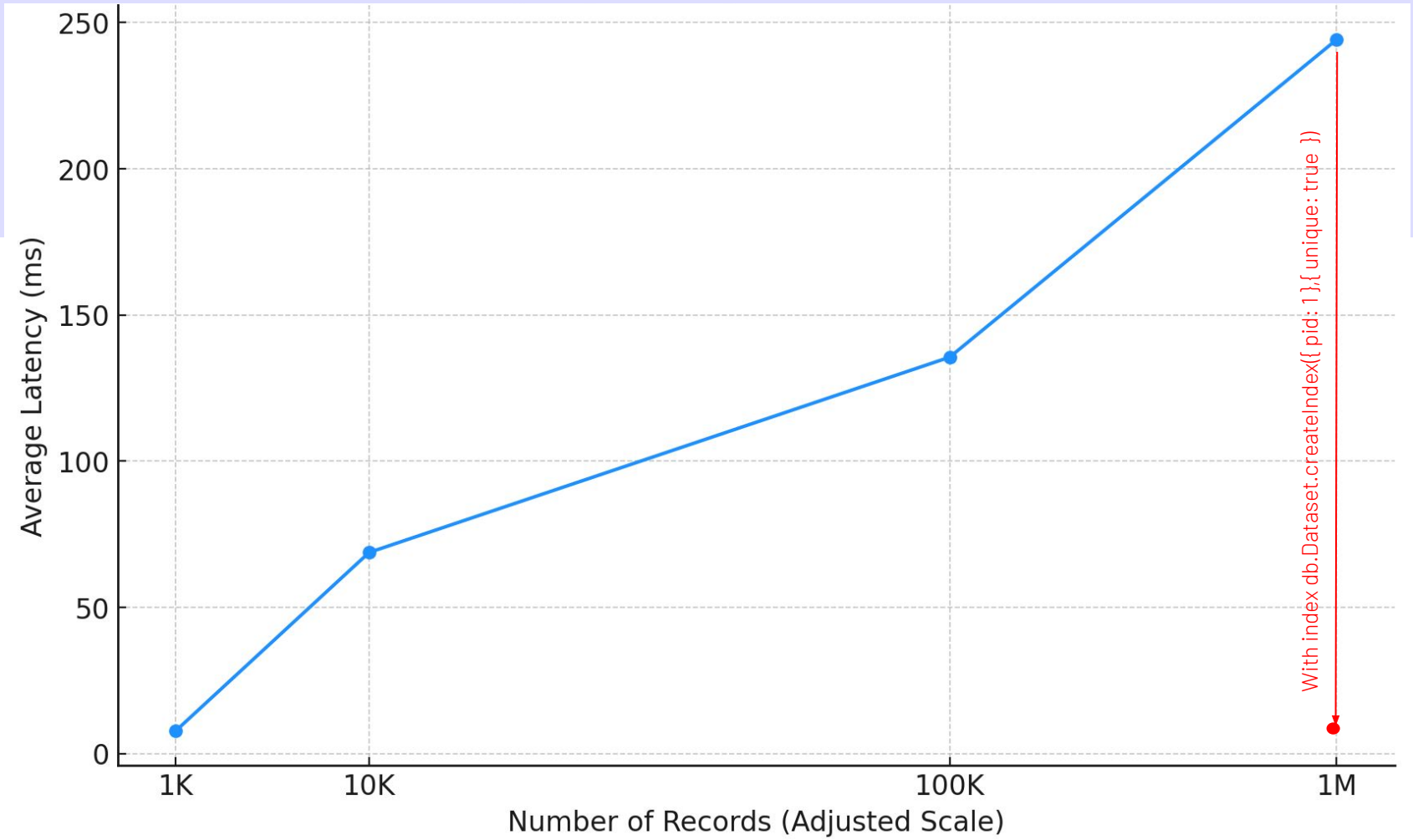
SciCat, part 1
Data-Out, PID

SciCatCon 2025

Igor Khokhriakov aka Ingvord

QUIZ

What result do you expect?


Average Latency VS Number of Records in the DB (C10, R1000)
Average Latency (ms)
250
200
150
100
50
0
1K
10K
100K
1M
Number of Records (Adjusted Scale)
With index db.Dataset.createIndex({ pid: 1},{ unique: true })

METHODOLOGY

Testing Tool:

- The benchmarking was conducted using wrk, a modern HTTP benchmarking tool capable of generating significant load with multithreading and connection options.

Test Scenarios:

- Multiple configurations were tested by varying:
  - **Dataset size**: 1K, 10K, 100K and 1M records.
  - **Request rate**: R10, R100, and R1000 requests per second.
  - **Concurrent connections**: 50 and 100 connections.

Endpoints:

- The REST API endpoint /datasets/fullquery was tested with query parameters to simulate real-world search patterns.

Metrics Collected:

- **Latency**: Average, standard deviation, and percentiles (50th, 75th, 90th, 99th).
- **Throughput**: Achieved requests per second and total requests processed.
- **Reliability**: Number of socket errors (timeouts, connection issues).

1_000 Records, 10 Clients, Rate {10..1000}

Configuration	Avg Latency	50th %ile Latency	99th %ile Latency	Requests/sec	Timeouts
R1000	9.83ms	4.78ms	163.97ms	1000.01	-
R100	12.06ms	10.81ms	30.91ms	100.05	-
R10	12.64ms	11.93ms	29.22ms	10.05	-

Conclusions

- **With 1K records, the system is highly efficient at all request rates**, even at **R1000, where it maintains excellent performance**.
- **Limiting clients to 10 (-c10) ensures stable performance without significant latency spikes**, even under high throughput.
- **Compared to 10K and 100K records, 1K datasets show minimal impact from increasing request rates**, indicating **the backend can efficiently handle smaller datasets at scale**.

10_000 Records, 10 Clients, Rate {10..1000}

Configuration	Avg Latency	50th %ile Latency	99th %ile Latency	Requests/sec	Timeouts
R1000	51.23ms	7.22ms	1.13s	999.73	-
R100	20.12ms	15.59ms	179.33ms	100.02	-
R10	16.80ms	16.42ms	36.13ms	10.06	-

Conclusions

- **Limiting clients to 10 results in significantly improved system performance** at all load levels.
- **Unlike 100K records, where R1000 introduced substantial latency spikes, 10K records show no major degradation even at R1000**.
- **The system scales well with 10K records**, achieving the target request rate with low latency.
- **For real-world usage, keeping concurrency controlled (e.g., -c10) can prevent unnecessary queuing delays and improve stability**.

100_000 Records, 10 Clients, Rate {10..1000}

Configuration	Avg Latency	50th %ile Latency	99th %ile Latency	Requests/sec	Timeouts
R1000	4.36s	2.91s	23.84s	**971.41**	-
R100	93.13ms	12.64ms	2.59s	**98.53**	-
R10	23.06ms	13.51ms	241.54ms	**10.02**	-

Conclusions

- **Lowering concurrent clients (-c10) significantly reduces queuing effects and improves system stability**.
- The system **scales well at low to moderate request rates (R10, R100)** but **still struggles under R1000**, with **high average and 99th percentile latencies**.
- **Compared to previous high concurrency tests, reducing client count has a major positive impact on performance**.

1_000_000 Records, 10 Clients, Rate {10..1000}

Configuration	Avg Latency	50th %ile Latency	99th %ile Latency	Requests/sec	Timeouts
R1000	127.60ms	6.82ms	2.08s	997.52	-
R100	14.49ms	11.60ms	74.30ms	100.06	-
R10	17.87ms	15.08ms	148.48ms	10.02	-

Conclusions

- **Lowering concurrent clients (-c10) significantly reduces queuing effects and improves system stability**.
- The system **scales well at low to moderate request rates (R10, R100)** but **still struggles under R1000**, with **high average and 99th percentile latencies**.
- **Compared to previous high concurrency tests, reducing client count has a major positive impact on performance**.

Combined Results Tableview

OLAP Cube - Read Performance Summary (Updated)

	Number of Records	50th Percentile Latenc	50th Percentile Latenc	50th Percentile Latenc	99th Percentile Latenc	99th Percentile Latenc	99th Percentile Latenc	Requests/sec (R10)	Requests/sec (R100)	Requests/sec (R1000)
1	1K	13.51	15.59	4.78	241.54	179.33	163.97	10.39	100.02	1000.01
2	10K	16.42	10.81	6.82	36.13	30.91	2080.0	10.02	100.05	997.52
3	100K	11.93	11.6	2910.0	29.22	74.3	23840.0	10.05	100.06	971.41
4	1M	15.08	12.64	7.22	148.48	2590.0	1130.0	10.02	98.53	999.73

CONCLUSIONS

- **System scales well under high request rates**, with **1K, 10K, 100K, and 1M records maintaining low latency at R1000**.
- **100K records initially showed a performance spike, but this was due to system overload (100% RAM and swap usage)**—when properly tested, **it performed reasonably well**.
- **Limiting concurrency (-c10) significantly improves stability**, reducing queuing effects and keeping latency distribution tighter.
- **1M records perform better than expected**, suggesting **efficient caching, indexing, or database optimizations at scale**.
- **Low (R10) and moderate (R100) request rates show excellent stability**, with **99th percentile latency remaining within acceptable ranges**.
- **At high load (R1000), 99th percentile latencies increase but remain within operational limits (~2s for 1M records)**.
- **Final takeaway**: The system handles high request loads efficiently across dataset sizes, and the **previous 100K anomaly serves as an anchor for discussions on system health monitoring**.

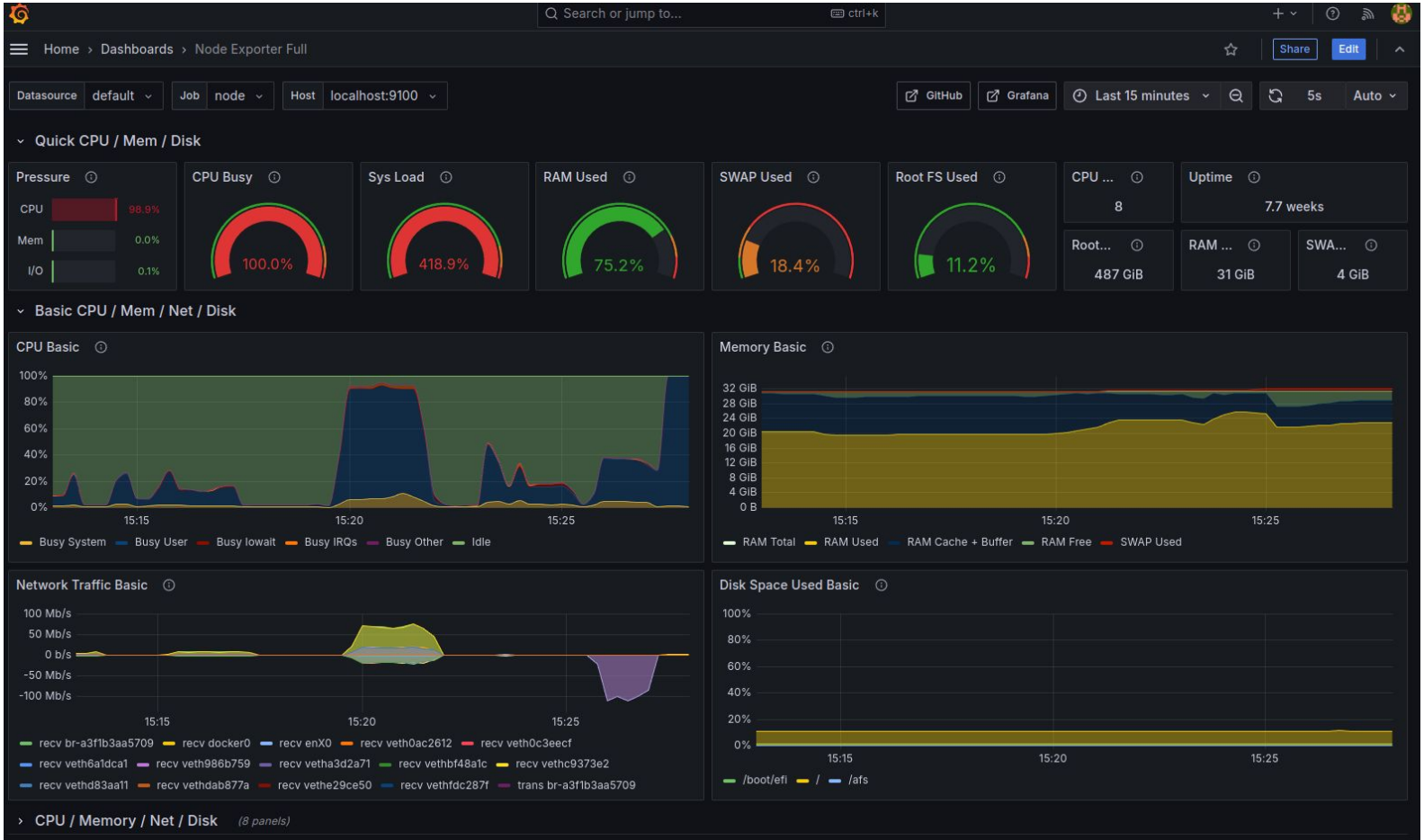
RECOMMENDATIONS

- Optimize database indexing and query execution plans for large datasets.
- Implement caching mechanisms for frequently accessed data.
- Conduct capacity planning to identify and address resource bottlenecks.
- Consider load balancing strategies to handle higher concurrency and throughput.
- Finally be ready to auto-scale the deployment.

SIDE NOTES

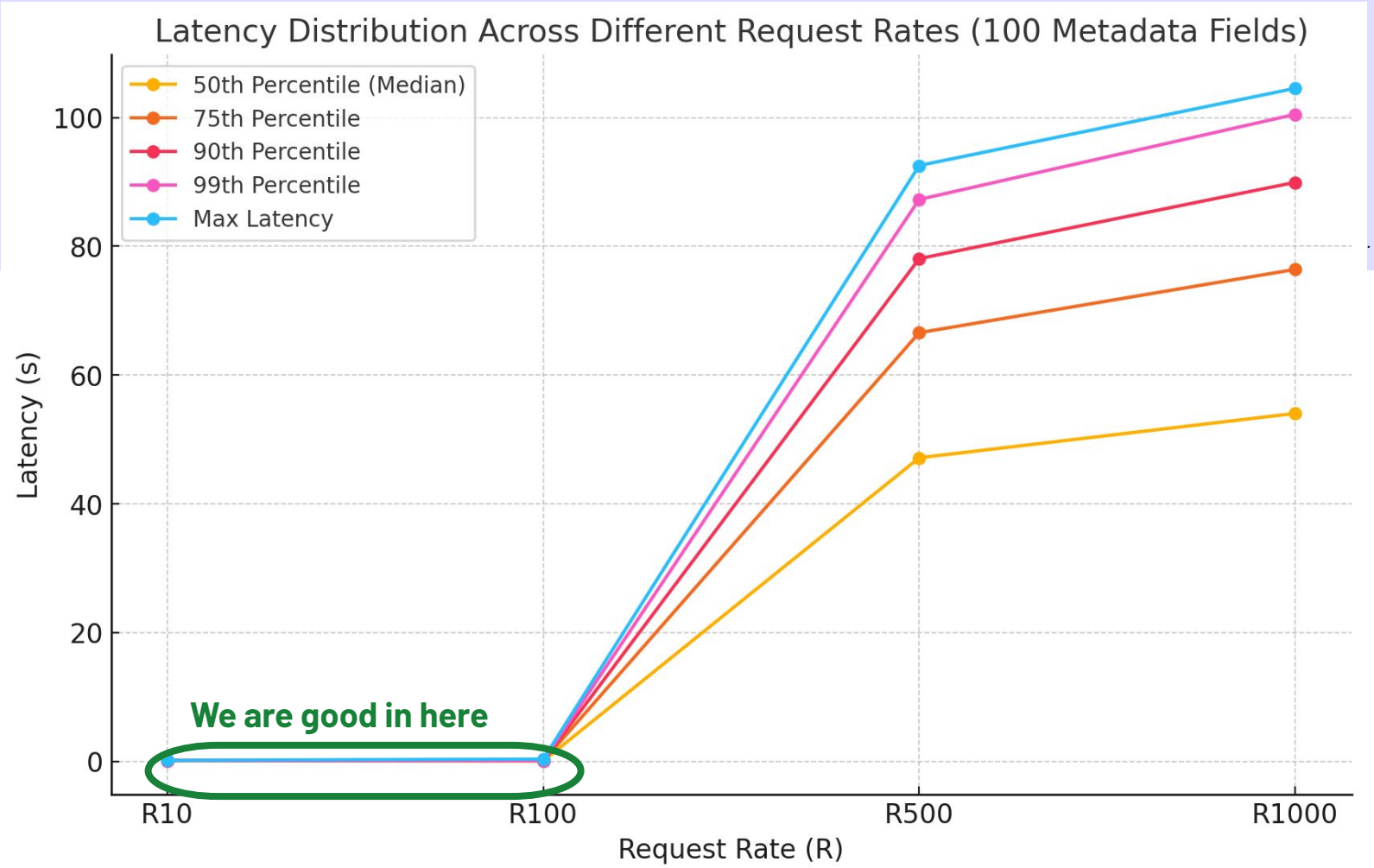

Search or jump to... ctrl+k
Home › Dashboards › Node Exporter Full
Share
Edit
Datasource default
Job node
Host localhost:9100
GitHub
Grafana
Last 15 minutes
5s
Auto
Quick CPU / Mem / Disk
Pressure
CPU 1.0%
Mem 0.1%
I/O 0.1%
CPU Busy
4.4%
Sys Load
30.6%
RAM Used
97.7%
SWAP Used
62.8%
Root FS Used
11.8%
CPU ...
8
Uptime
1.7 days
Root...
487 GiB
RAM ...
31 GiB
SWA...
4 GiB
Basic CPU / Mem / Net / Disk
CPU Basic
100%
80%
60%
40%
20%
0%
10:45
10:50
10:55
Busy System
Busy User
Busy Iowait
Busy IRQs
Busy Other
Idle
Memory Basic
36 GiB
32 GiB
28 GiB
24 GiB
20 GiB
16 GiB
12 GiB
8 GiB
4 GiB
0 B
10:45
10:50
10:55
RAM Total
RAM Used
RAM Cache + Buffer
RAM Free
SWAP Used
Network Traffic Basic
20 Mb/s
15 Mb/s
10 Mb/s
5 Mb/s
0 b/s
-5 Mb/s
10:45
10:50
10:55
recv br-a3f1b3aa5709
recv docker0
recv enX0
recv veth41013f9
recv veth4a5e294
recv veth6277448
recv veth737ce9c
recv veth783de17
recv veth7ef6a05
recv veth7ffcc7c
recv vethc959d61
recv vethcf3ad63
recv vethf2884f5
trans br-a3f1b3aa5709
trans docker0
Disk Space Used Basic
100%
80%
60%
40%
20%
0%
10:45
10:50
10:55
/boot/efi
/
/afs
CPU / Memory / Net / Disk
(8 panels)

SYSTEM LOAD 100_000, c=100, R=1000



SciCat, part 2
Data-In, Metadata

Igor Khokhriakov aka Ingvord


Latency Distribution Across Different Request Rates (100 Metadata Fields)
50th Percentile (Median)
75th Percentile
90th Percentile
99th Percentile
Max Latency
Latency (s)
0
20
40
60
80
100
We are good in here
R10
R100
R500
R1000
Request Rate (R)

CONCLUSIONS

Summary of Ingestion Performance Tests (100 Metadata Fields)

✅ **Stable performance at low request rates (`R10`, `R100`)**

- Low latency (**50th percentile: 86ms at `R10`, 65ms at `R100`**).
- The system handles these loads **efficiently without degradation**.

❌ **Severe performance degradation at higher loads (`R500`, `R1000`)**

- **50th percentile latency jumps to 47s (`R500`) and 54s (`R1000`)**.
- **99th percentile reaches 87s (`R500`) and over 100s (`R1000`)**, making real-time ingestion infeasible.

❌ **Backend reaches a hard capacity limit (~140-150 req/sec)**

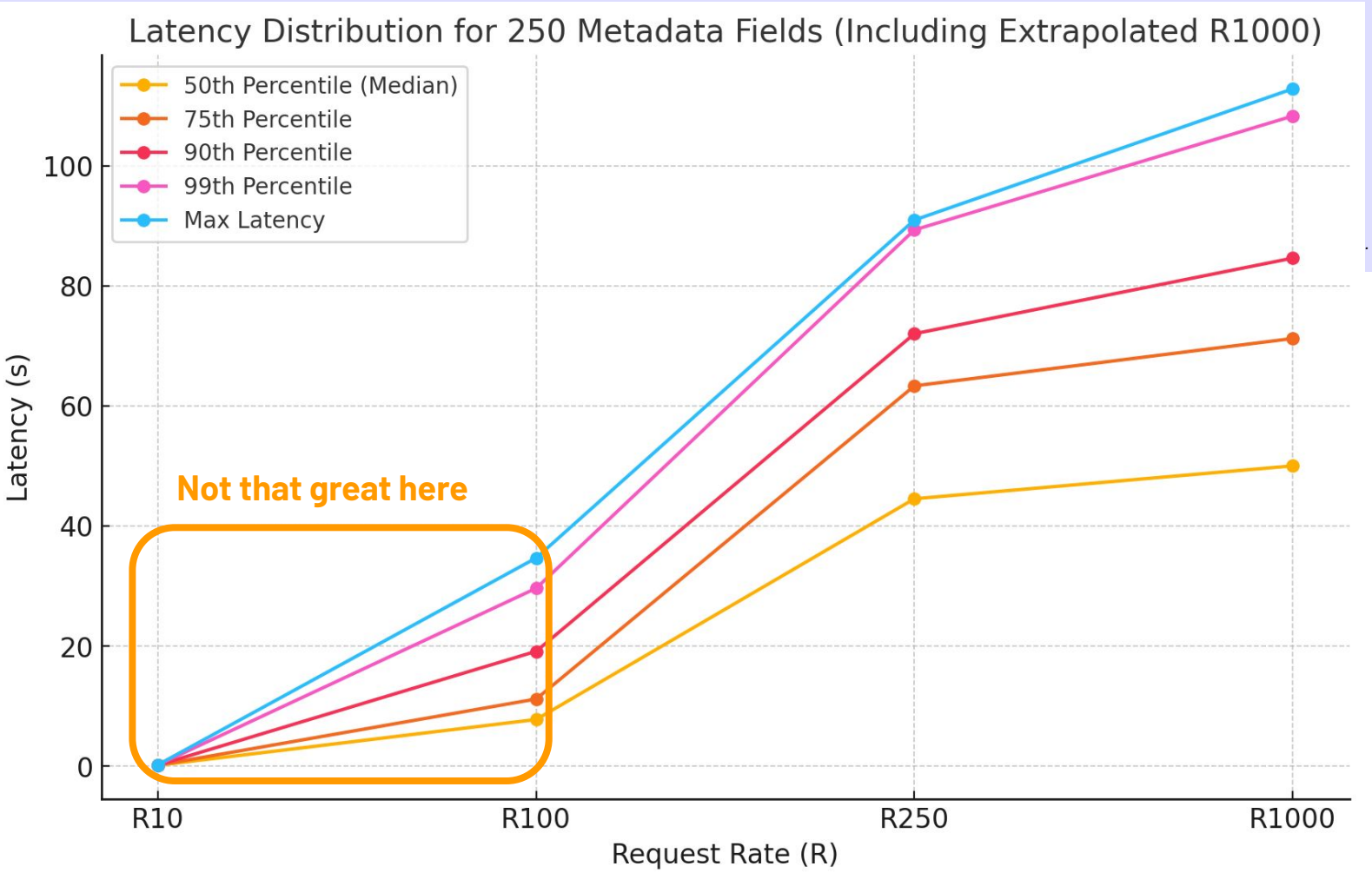
- At `R500` and `R1000`, the system **only achieves ~140-150 req/sec**, far below the target rates.
- Indicates **queueing, serialization, or database bottlenecks**.

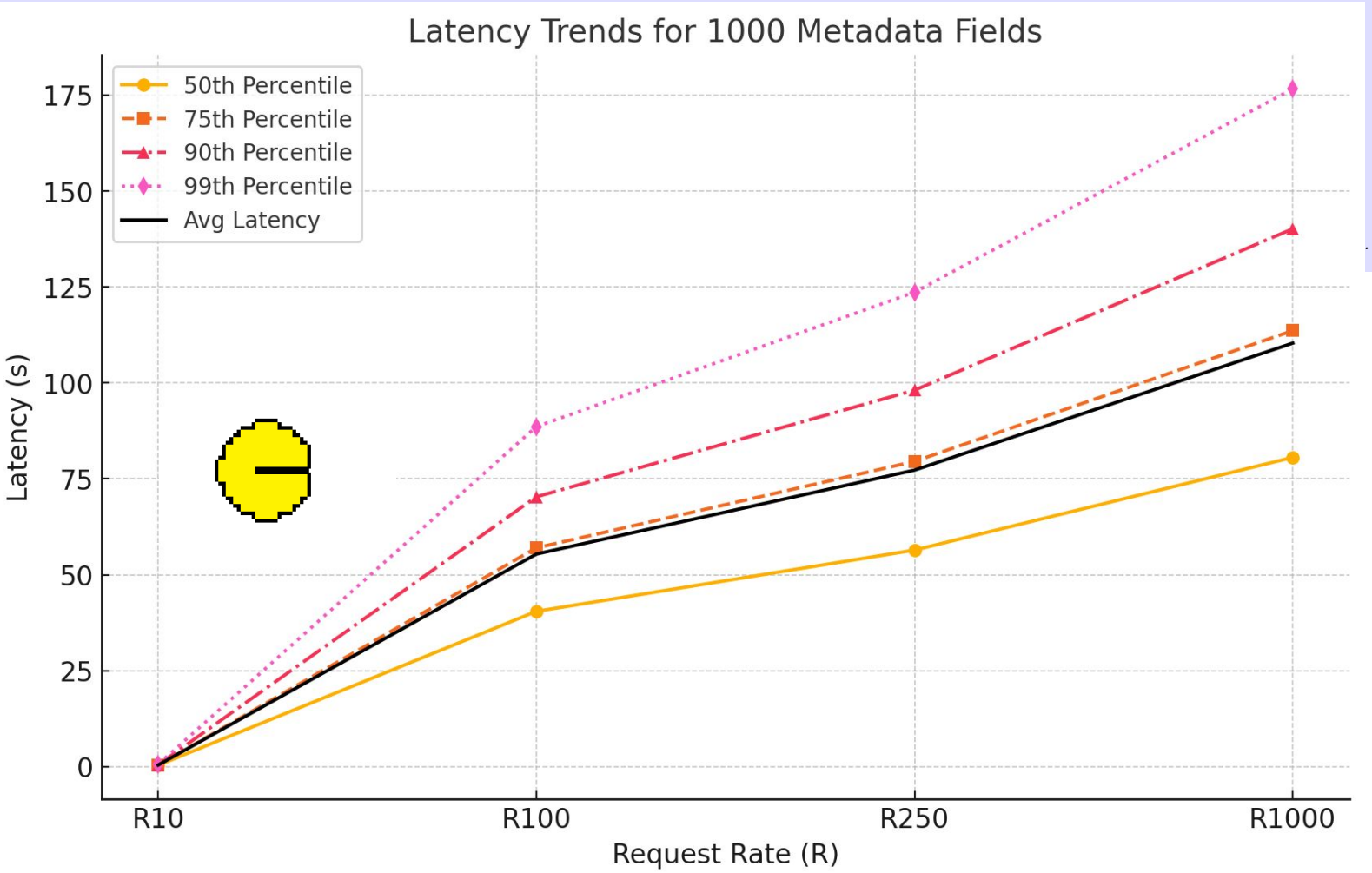
✅ **No outright failures, but long processing delays**

- No **timeouts or rejected requests**, meaning the system is **processing all requests but at a slow pace**.

❌ **Lack of horizontal scalability observed**

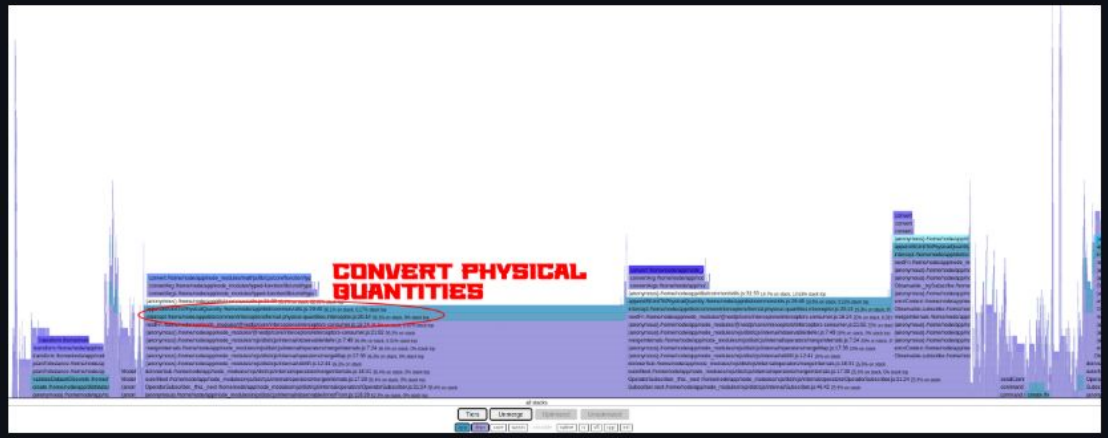
- **Higher concurrency does not improve throughput**.
- Suggests **database, transaction, or resource contention issues**.


Latency Distribution for 250 Metadata Fields (Including Extrapolated R1000)
50th Percentile (Median)
75th Percentile
90th Percentile
99th Percentile
Max Latency
Latency (s)
0
20
40
60
80
100
Not that great here
R10
R100
R250
R1000
Request Rate (R)


Latency Trends for 1000 Metadata Fields
50th Percentile
75th Percentile
90th Percentile
99th Percentile
Avg Latency
Latency (s)
175
150
125
100
75
50
25
0
R10
R100
R250
R1000
Request Rate (R)

Ingvord on Mar 23 Member Author ···

Jo, I pinned down the suspect - it is `convert-physical-quantities-interceptor`. I this in the flame grpah:

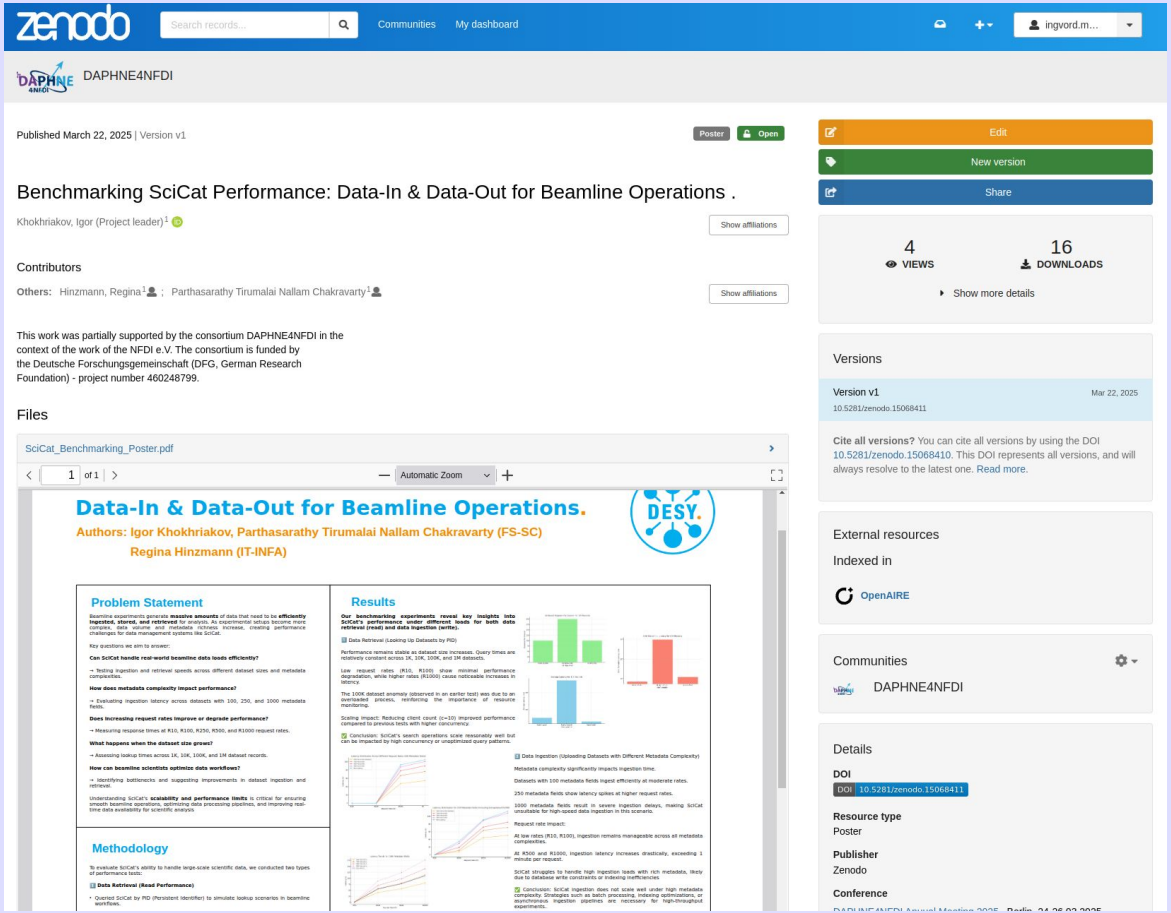


As well as in APM:

tetur adipiscing elit.

MORE DETAILS IN THE DAPHNE4NFDI ANNUAL MEETING POSTER

https://zenodo.org/records/15068411



SciCat, part 3
Data-In, Origdatablocks

Igor Khokhriakov aka Ingvord

OrigDatablocks

Uploading 1M DataFiles

```
Response Status: 413
Body: <html>
<head><title>413 Request Entity Too Large</title></head>
<body>
<center><h1>413 Request Entity Too Large</h1></center>
<hr><center>nginx/1.24.0 (Ubuntu)</center>
</body>
</html>
```

```
# client buffers
client_max_body_size 100M;  # max allowed body size
client_body_buffer_size 10M;
client_header_buffer_size 16k;
# useful for big headers e.g. tokens, large queries etc
large_client_header_buffers 4 32k;
proxy_buffers 16 64K;
proxy_buffer_size 64k;
proxy_busy_buffers_size 128k;
proxy_read_timeout 1200;
proxy_send_timeout 60;

# prevent buffering large responses in memory
proxy_max_temp_file_size 0;

proxy_request_buffering off;
```

Error occurrence

3 minutes ago | POST http://localhost/backend/api/v3/origdatablocks | 500 Internal Server Error | POST unknown route

Exception stack trace | Metadata

request entity too large

```
› at readStream (node_modules/raw-body/index.js:163)
› at getRawBody (node_modules/raw-body/index.js:116)
› at read (node_modules/body-parser/lib/read.js:79)
› at jsonParser (node_modules/body-parser/lib/types/json.js:138)
› at handle (node_modules/elastic-apm-node/lib/instrumentation/modules/express.js:80)
› at handle (node_modules/express/lib/router/layer.js:95)
› at trim_prefix (node_modules/express/lib/router/index.js:328)
› at <anonymous> (node_modules/express/lib/router/index.js:286)
› at process_params (node_modules/express/lib/router/index.js:346)
› at next (node_modules/express/lib/router/index.js:280)
› at arguments.<computed> (node_modules/elastic-apm-node/lib/instrumentation/modules/express.js:76)
› at cors (node_modules/cors/lib/index.js:188)
› at <anonymous> (node_modules/cors/lib/index.js:224)
› at originCallback (node_modules/cors/lib/index.js:214)
› at <anonymous> (node_modules/cors/lib/index.js:219)
› at optionsCallback (node_modules/cors/lib/index.js:199)
› at corsMiddleware (node_modules/cors/lib/index.js:204)
› at handle (node_modules/elastic-apm-node/lib/instrumentation/modules/express.js:80)
› at handle (node_modules/express/lib/router/layer.js:95)
› at trim_prefix (node_modules/express/lib/router/index.js:328)
› at <anonymous> (node_modules/express/lib/router/index.js:286)
› at process_params (node_modules/express/lib/router/index.js:346)
› at next (node_modules/express/lib/router/index.js:280)
› at expressInit (node_modules/express/lib/middleware/init.js:40)
› at handle (node_modules/elastic-apm-node/lib/instrumentation/modules/express.js:80)
› at handle (node_modules/express/lib/router/layer.js:95)
› at trim_prefix (node_modules/express/lib/router/index.js:328)
› at <anonymous> (node_modules/express/lib/router/index.js:286)
› at process_params (node_modules/express/lib/router/index.js:346)
› at next (node_modules/express/lib/router/index.js:280)
› at query (node_modules/express/lib/middleware/query.js:45)
› at handle (node_modules/elastic-apm-node/lib/instrumentation/modules/express.js:80)
› at handle (node_modules/express/lib/router/layer.js:95)
› at trim_prefix (node_modules/express/lib/router/index.js:328)
› at <anonymous> (node_modules/express/lib/router/index.js:286)
```

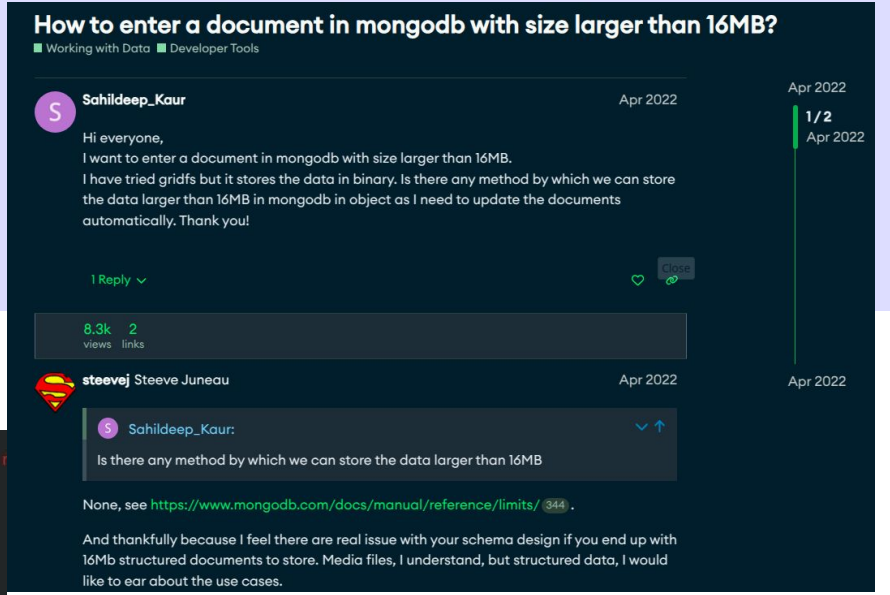
```
const config :  = {
  maxFileUploadSizeInMb: process.env.MAX_FILE_UPLOAD_SIZE || "16mb", // 16MB by default
```

OrigDatablocks

16MB Mongo document size limit

```
[Nest] 1  - 06/03/2025, 11:39:11 AM   ERROR "Internal server error"
[Nest] 1  - 06/03/2025, 11:39:11 AM   ERROR RangeError [ERR_OUT_OF_RANGE]: The value of "offset" is out of r
[Nest] 1  - 06/03/2025, 11:39:11 AM   ERROR Object:
{
  "requestUrl": "/backend/api/v3/origdatablocks",
  "requestUser": {
    "_id": "67d721c2948c9ce74a78d84a",
    "username": "admin",
    "email": "admin@scicat.project",
    "currentGroups": [
      "globalaccess",
      "admin"
    ],
    "authStrategy": "local"
  },
  "statusCode": 500
}
```

Number of DataFiles limit is ~75K



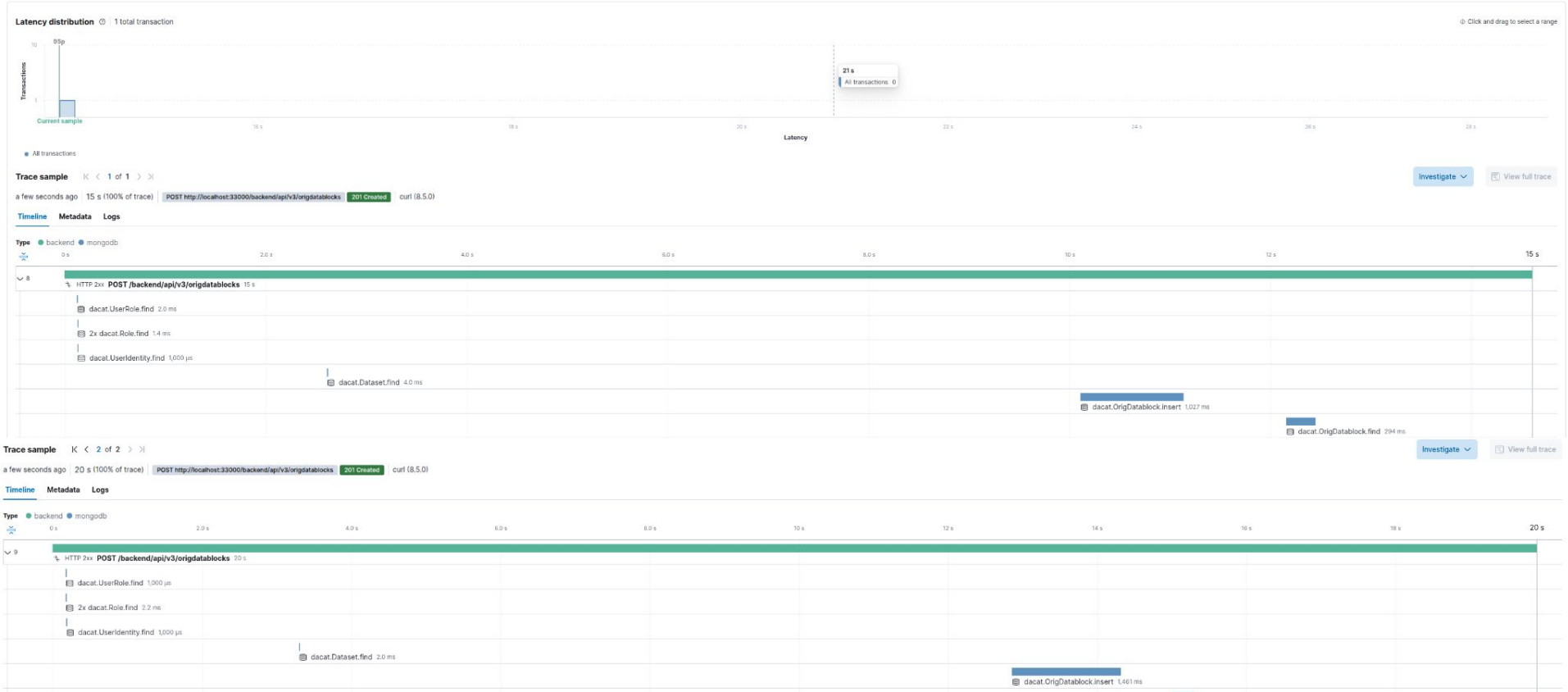
Can I increase document limit size in mongo collection?

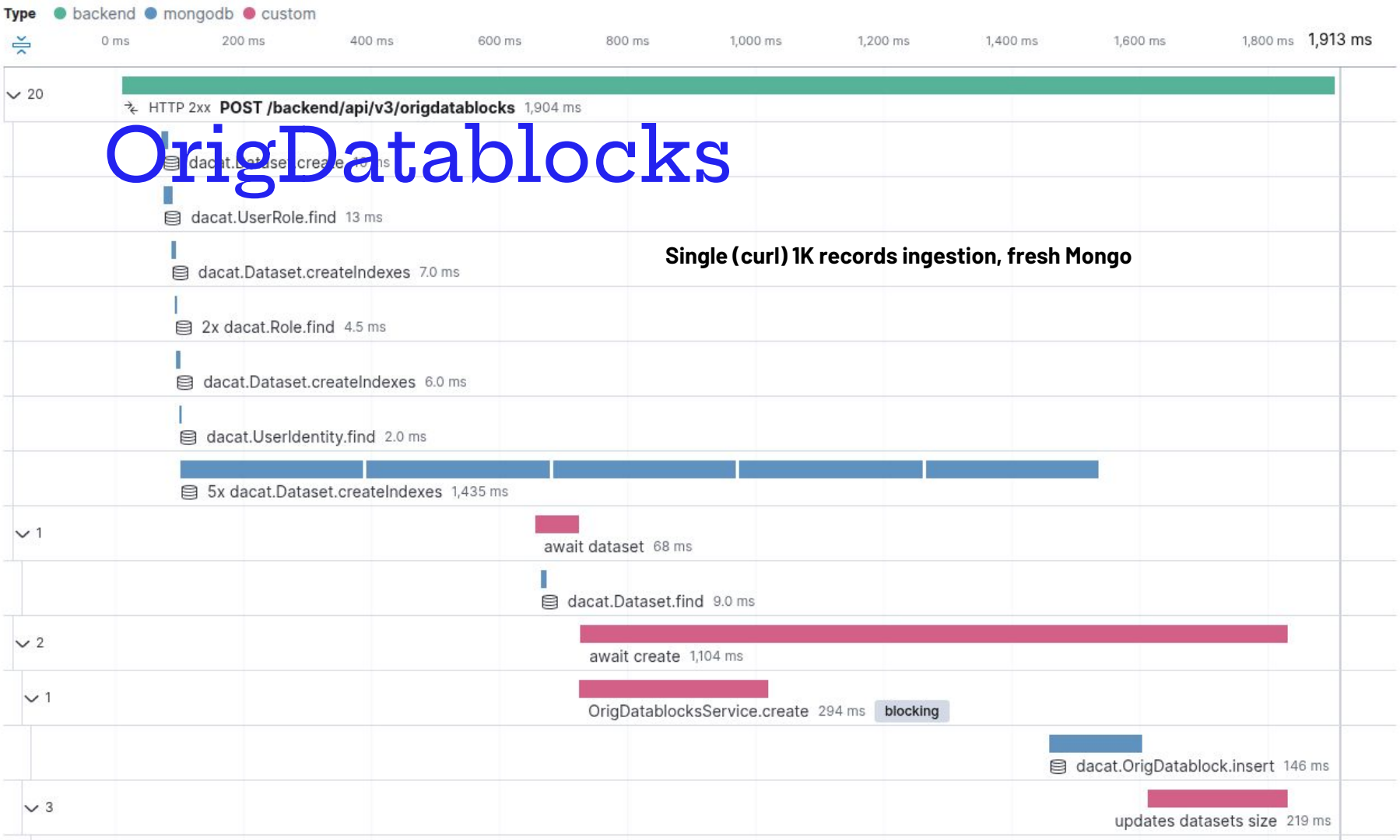
No, you **cannot increase the document size limit** in MongoDB — it's a **hard limit of 16MB** per document, and cannot be changed even through configuration or internal hacks.

🔒 Why?

- MongoDB is designed with a **16MB document limit** to ensure performance and memory efficiency.
- This limit applies to **any single BSON document**, including embedded arrays and subdocuments

OrigDatablocks

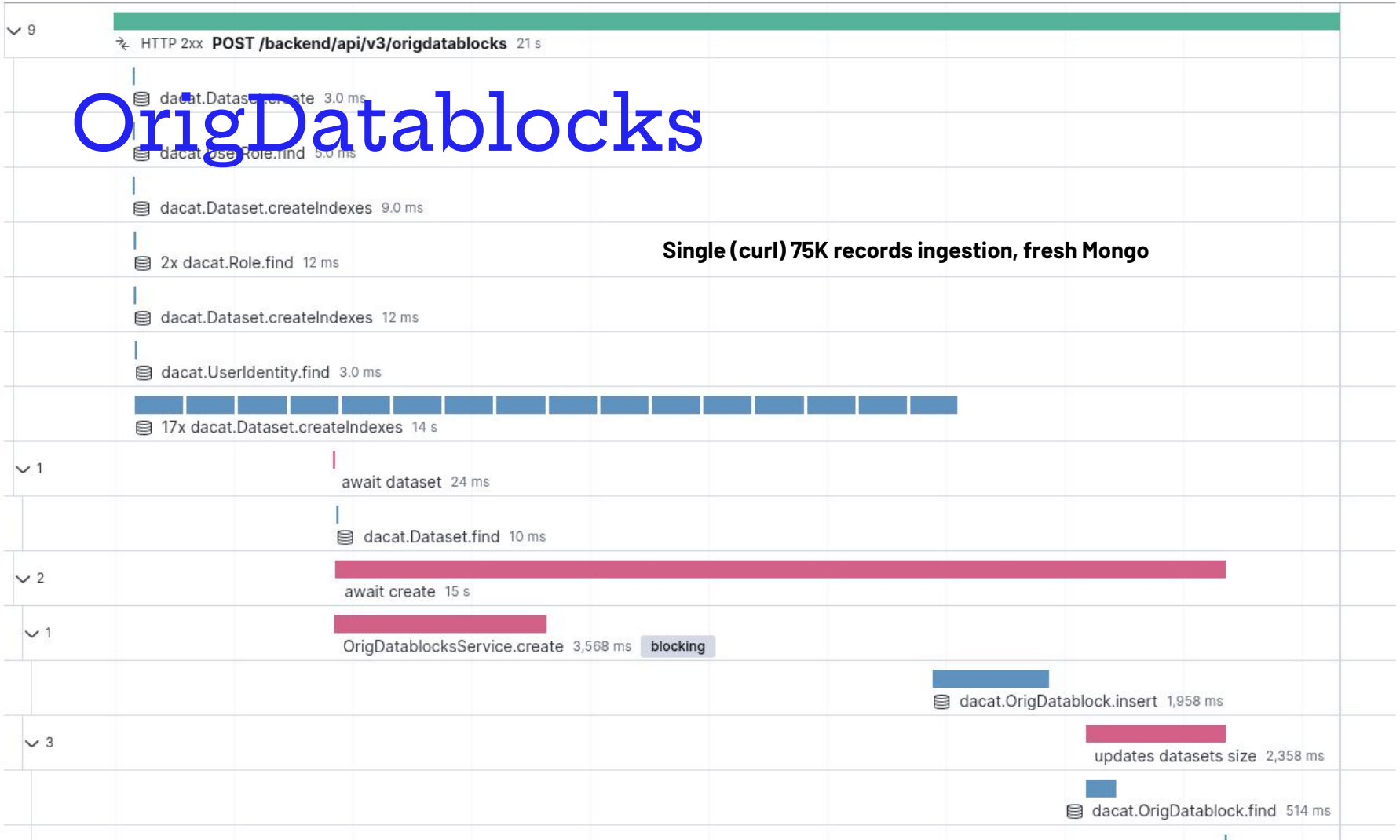




OrigDatablocks

Single (curl) 1K records ingestion, fresh Mongo

OrigDatablocks



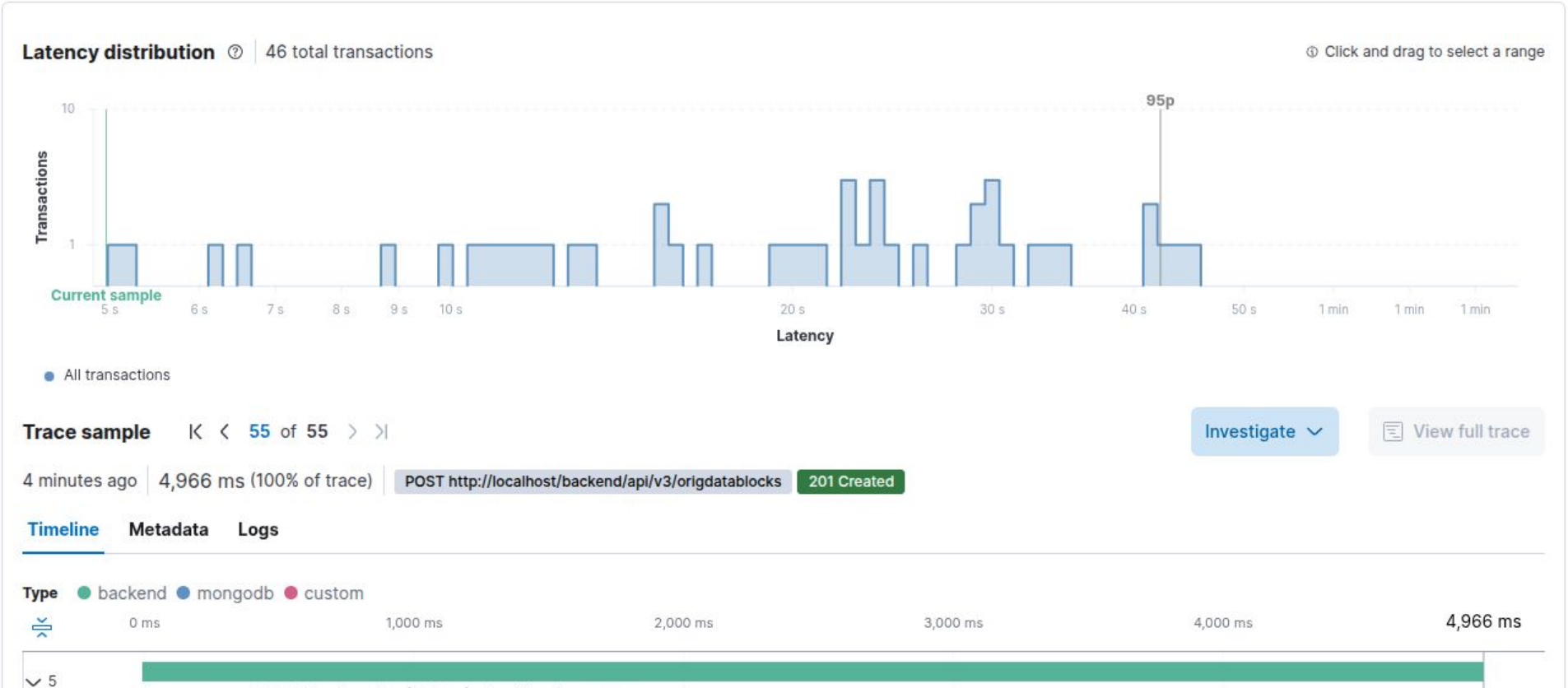
Single (curl) 75K records ingestion, fresh Mongo

OrigDatablocks

1K records ingestion (-c10 -t10 -R10)

OrigDatablocks

10K records ingestion (-c10 -t10 -R10)



OrigDatablocks

Trace sample 1 of 11

Investigate

View full trace

2 minutes ago | 86 s (100% of trace) | POST http://localhost/backend/api/v3/origdatablocks | 201 Created

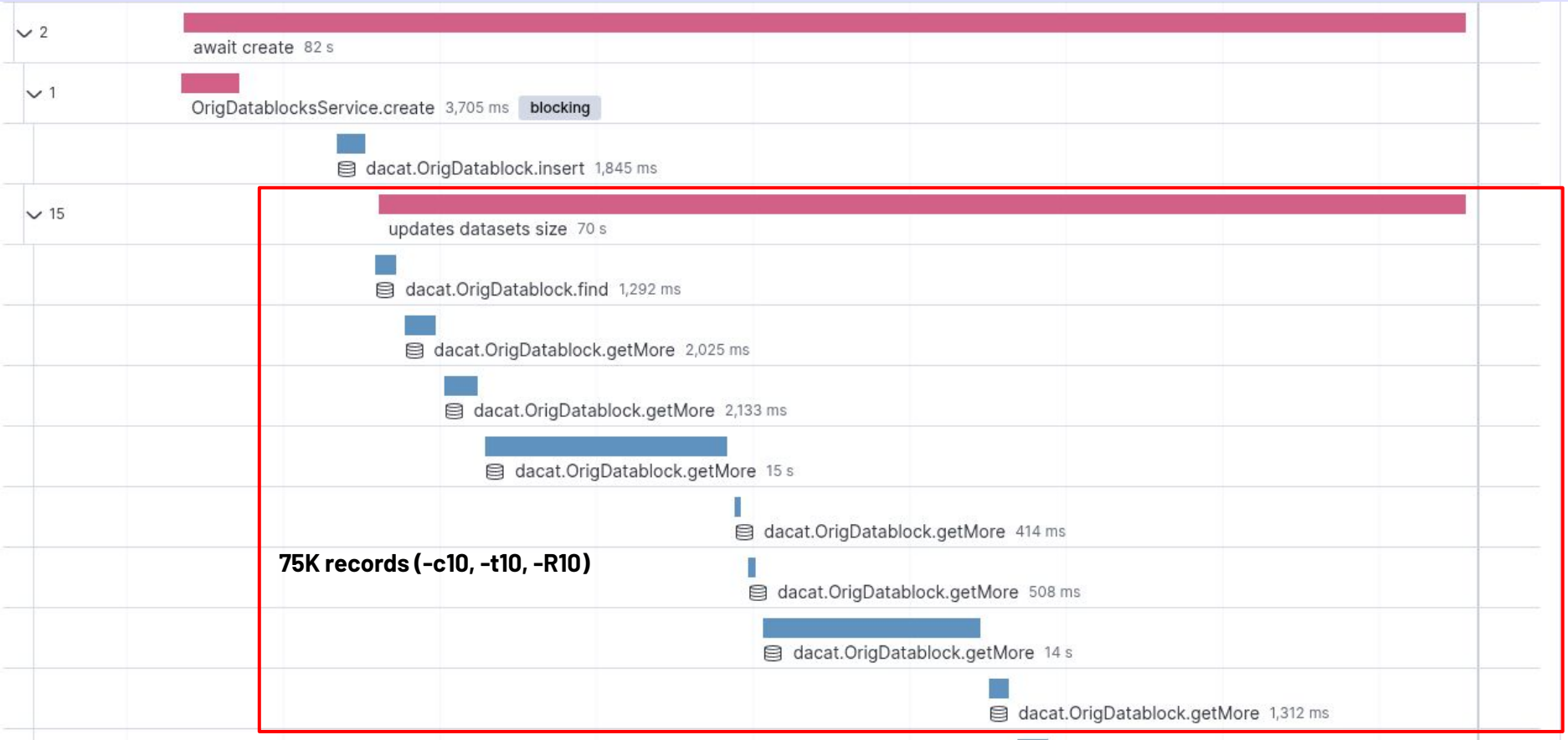
Timeline Metadata Logs

Type backend mongodb custom

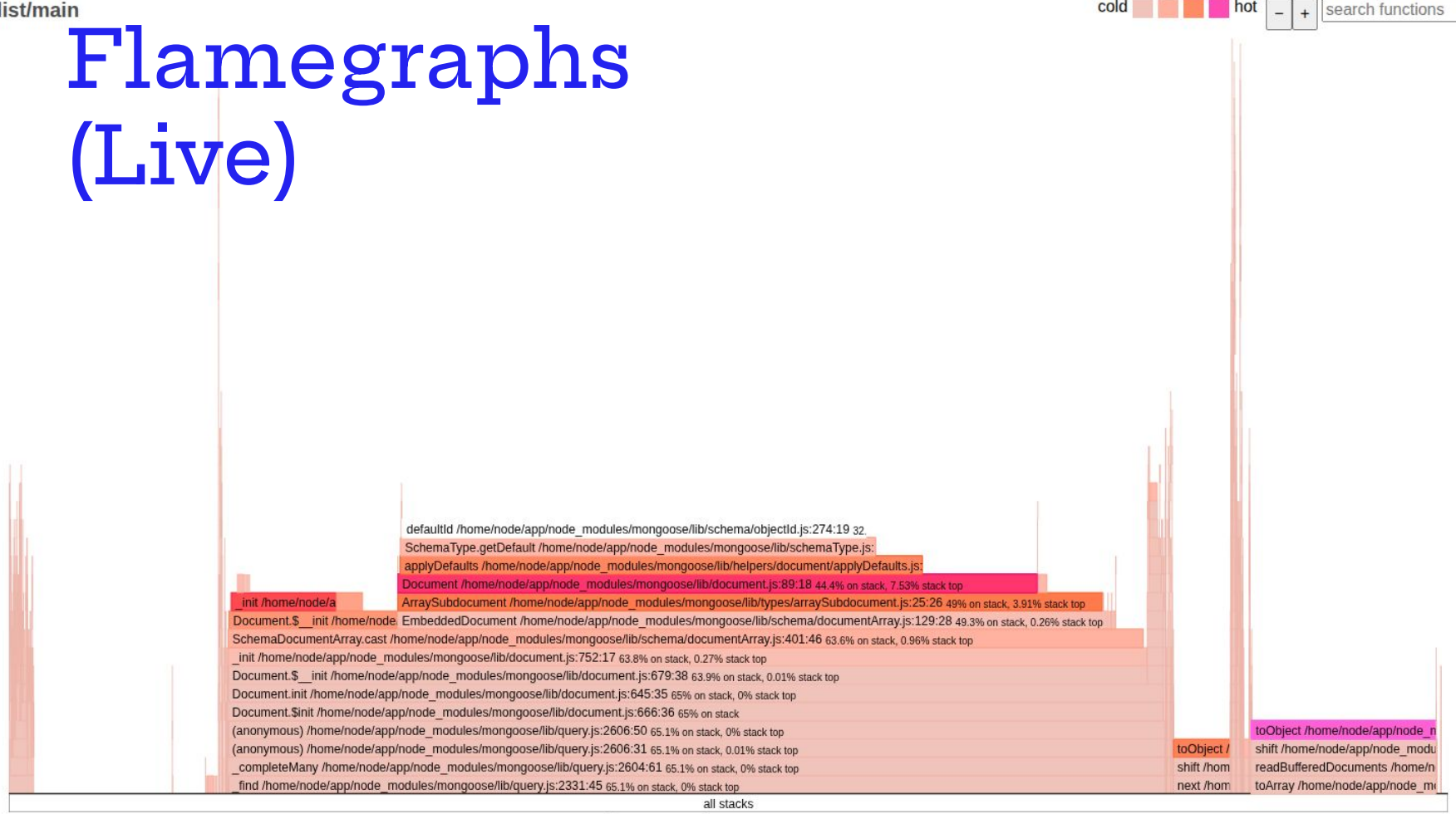
0 s 10 s 20 s 30 s 40 s 50 s 60 s 70 s 80 s 86 s

6 success POST /backend/api/v3/origdatablocks 86 s

OrigDatablocks



Flamegraphs (Live)



https://ingvord.github.io/animated-garbanzo

CONCLUSIONS


```
213   async updateDatasetSizeAndFiles(pid: string) : Promise<void>  {  Show usages  Max Novelli +1
214     // updates datasets size
215     const parsedFilters: IFilters<OrigDatablockDocument, IOrigDatablockFields> =
216       { where: { datasetId: pid } };
217     const datasetOrigdatablocks : OrigDatablock[]  =
218       await this.origDatablocksService.findAll(parsedFilters);
219
220     const updateDatasetDto: PartialUpdateDatasetDto = {
221       size: datasetOrigdatablocks
222         .map((od : OrigDatablock ) => od.size)
223         .reduce((ps : number , a : number ) => ps + a, 0),
224       numberOfFiles: datasetOrigdatablocks
225         .map((od : OrigDatablock ) => od.dataFileList.length)
226         .reduce((ps : number , a : number ) => ps + a, 0),
227     };
```

- **MongoDB `getMore` feeds batches of documents to the app (APM)**
- **Mongoose applies defaults + initializes models for each document fetched (flamegraph)**
- **Together: high latency + CPU cost arise from fetching + transforming large document sets**
- **Optimization target: use aggregate or lean queries (https://mongoosejs.com/docs/tutorials/lean.html)**

AVAILABLE FOR HIRE STARTING IN SEP'25

Igor Khokhriakov

Hamburg, Hamburg, Germany · **Contact Info**

3K followers · 500+ connections

See your mutual connections

DESY

Russian State University of Humanities

Websites

K8S realm
4 replicas
2 mongoDb rs

Igor Khokhriakov aka Ingvord

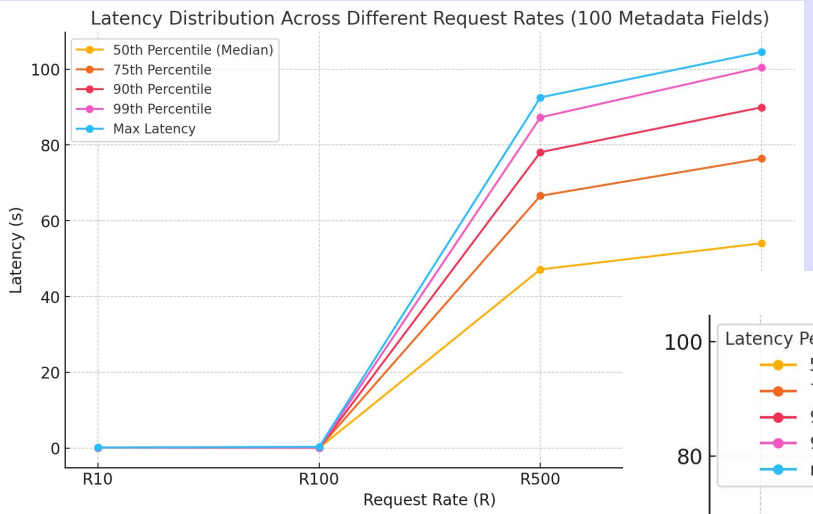
1_000_000 Records, 10 Clients, Duration 120s, Rate {10..1000}

Local VM VS K8s setup
Data-out, PID (w/index)

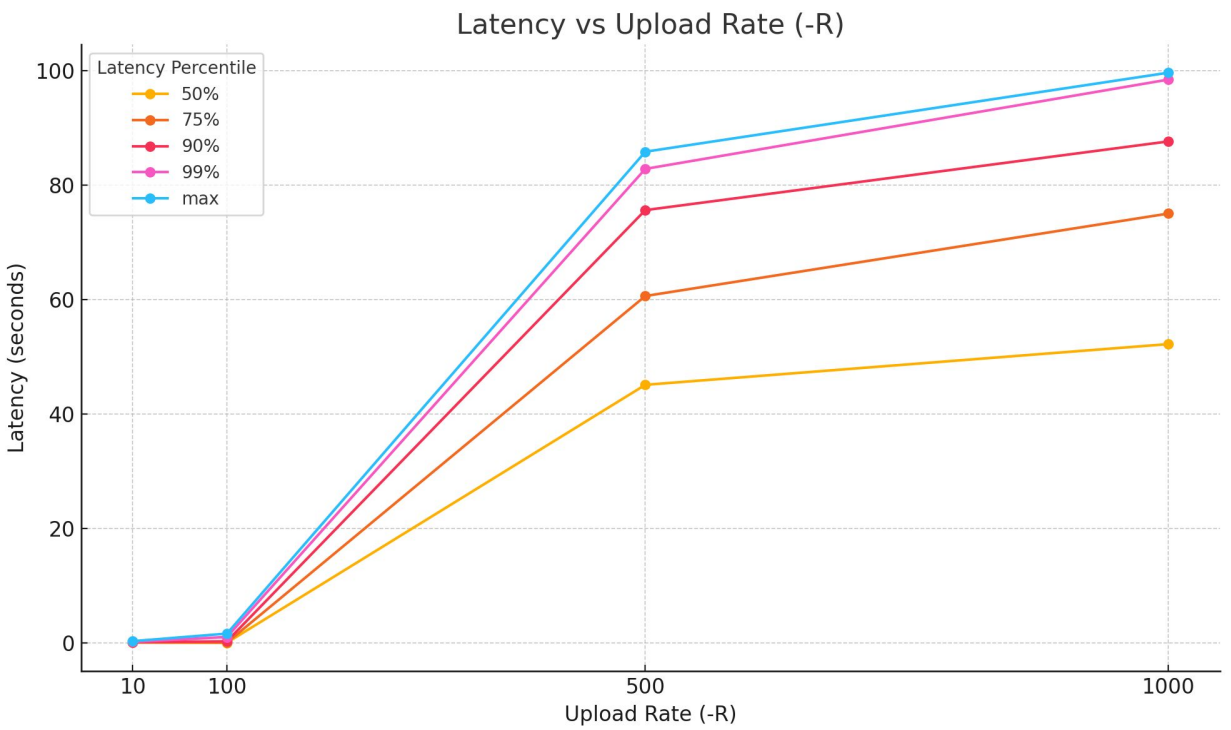
Configuration	Avg Latency	50th %ile Latency	99th %ile Latency	Requests/sec	Timeouts
R1000	127.60ms	6.82ms	2.08s	997.52	-
R100	14.49ms	11.60ms	74.30ms	100.06	-
R10	17.87ms	15.08ms	148.48ms	10.02	-

-R	Total Requests	Avg Latency	50% Latency	99% Latency	Max Latency	Req/sec
1000	120000	47.72ms	6.47ms	786.94ms	1.03s	999.97
100	12010	9.01ms	8.60ms	15.44ms	29.90ms	100.07
10	1209	15.96ms	12.90ms	32.96ms	70.85ms	10.07

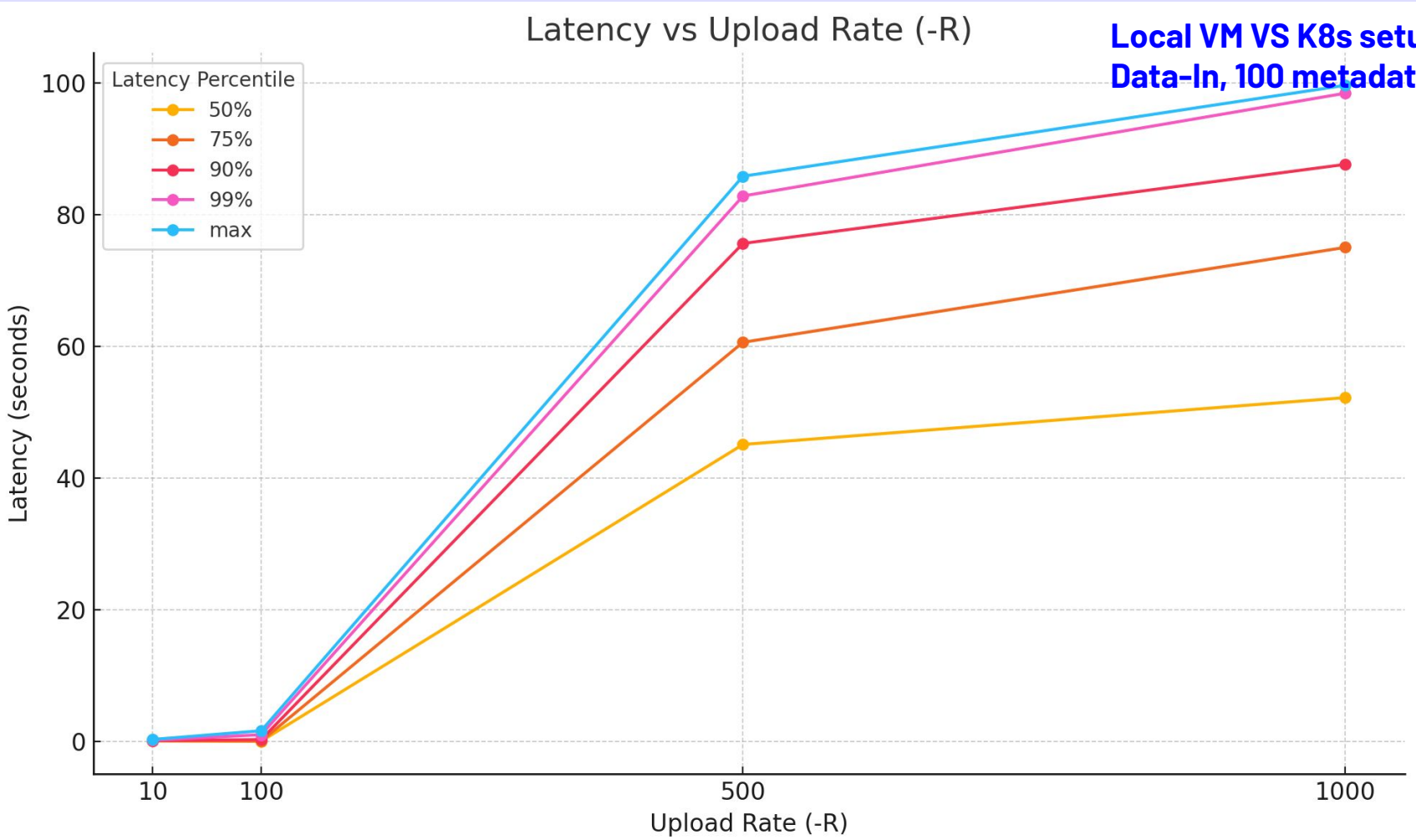
K8s setup significantly better at 99% percentile (*mongoDb rs*?!)

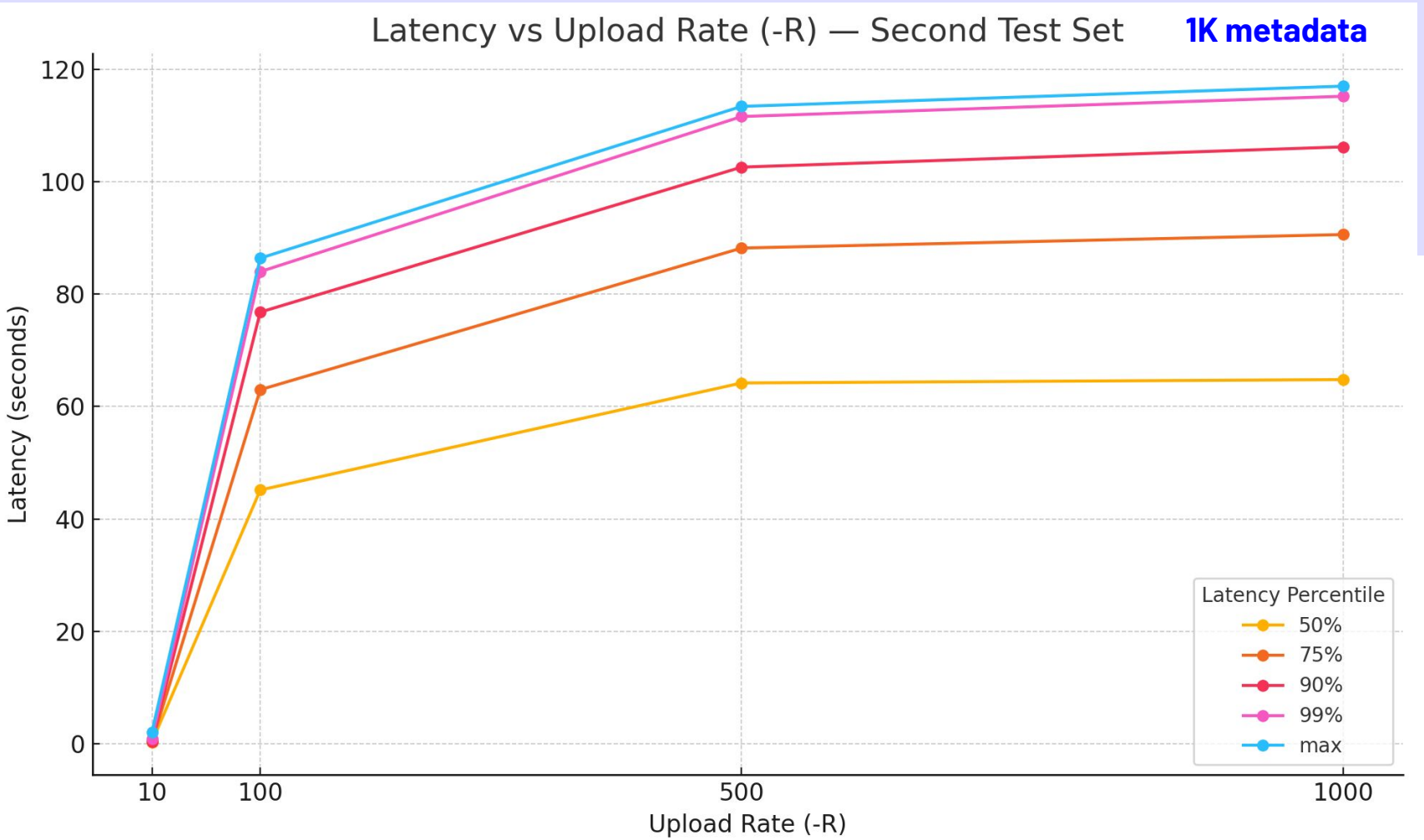


Local VM VS K8s setup
Data-In, 100 metadata

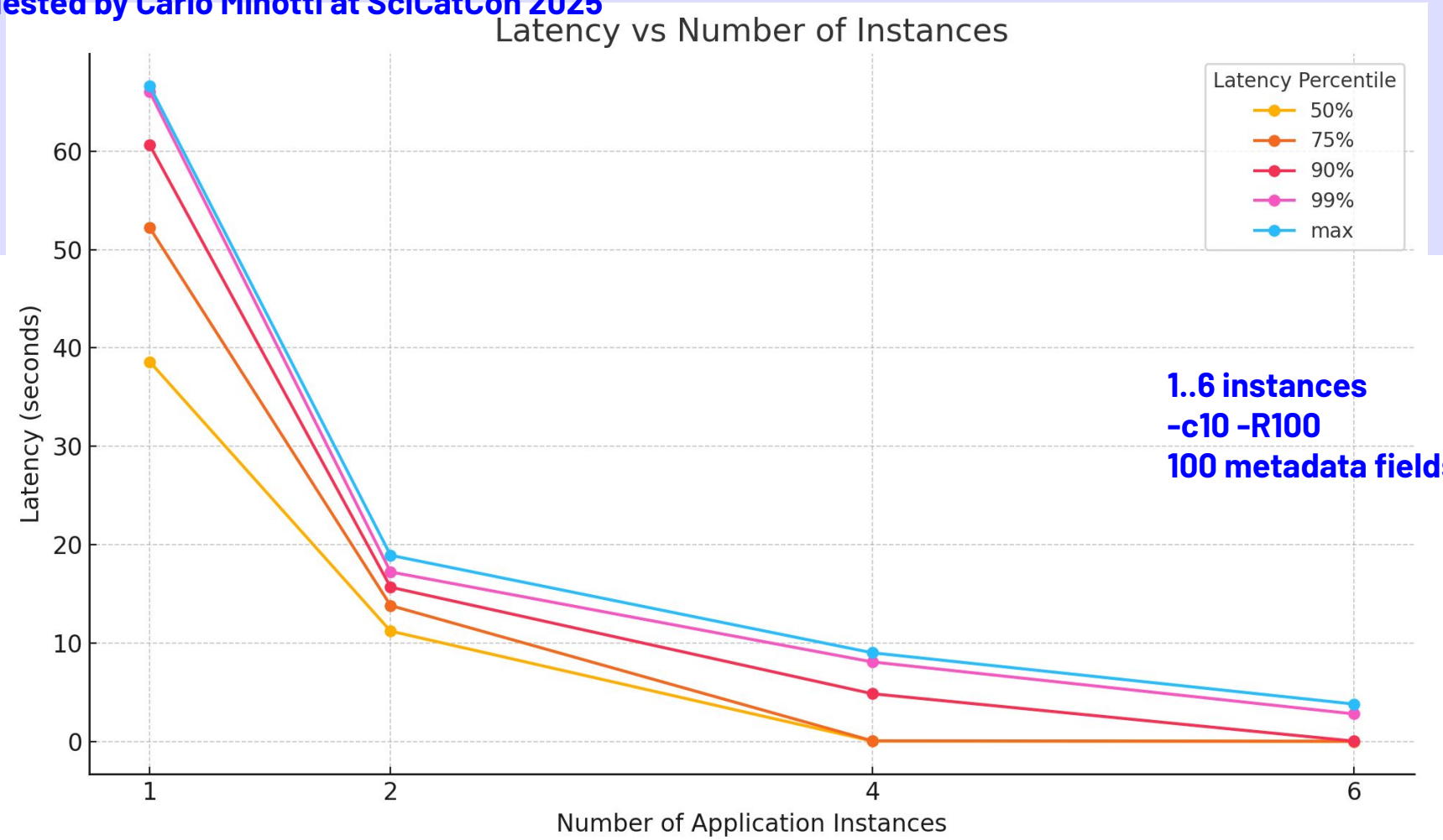


Yes, mostly identical
(*as expected - due to application code*)


Latency vs Upload Rate (-R)
Local VM VS K8s setup
Data-In, 100 metadata
Latency Percentile
50%
75%
90%
99%
max
Latency (seconds)
0
20
40
60
80
100
10
100
500
1000
Upload Rate (-R)

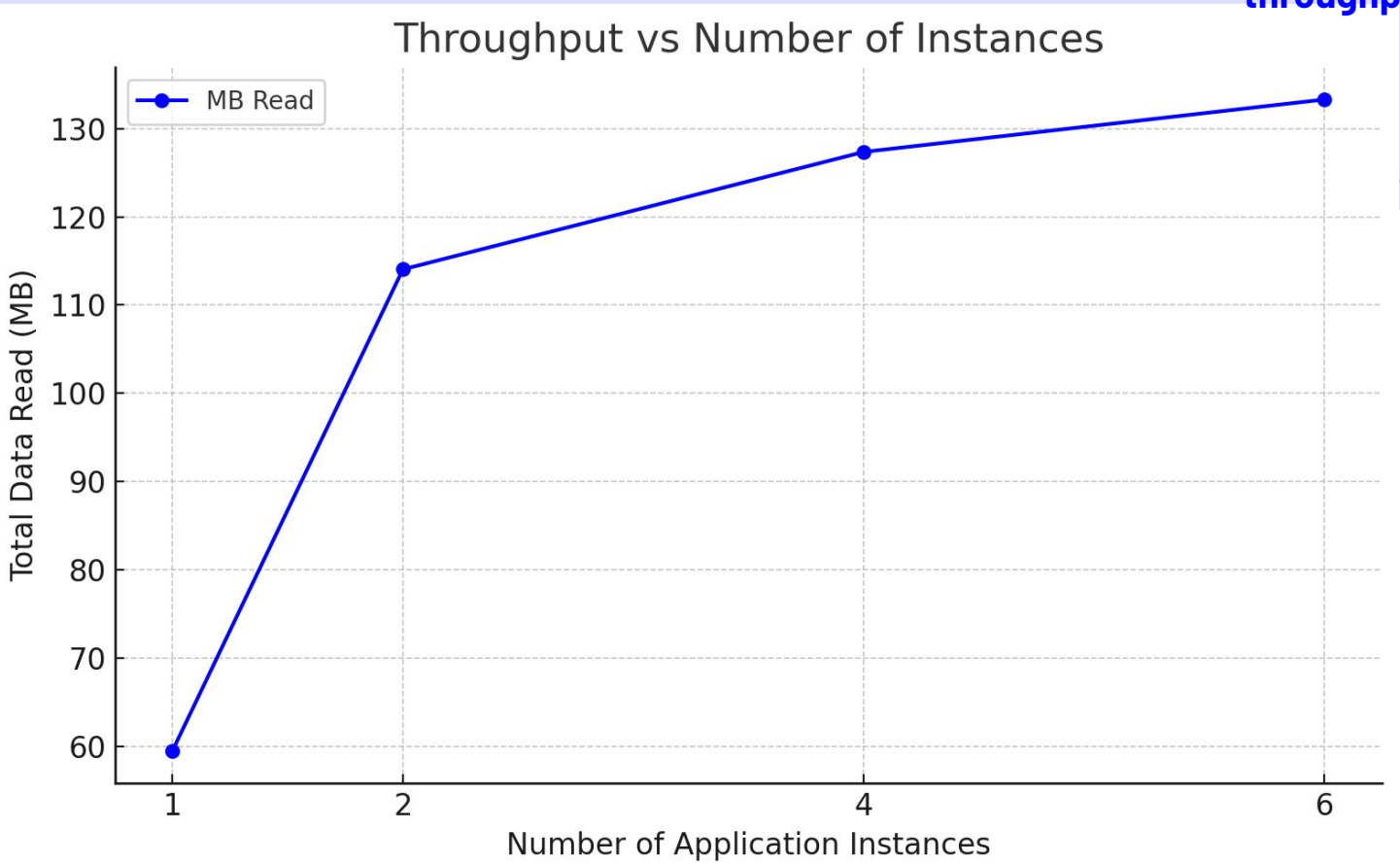

Latency vs Upload Rate (-R) — Second Test Set
1K metadata
Latency (seconds)
Upload Rate (-R)
0
20
40
60
80
100
120
10
100
500
1000
Latency Percentile
50%
75%
90%
99%
max

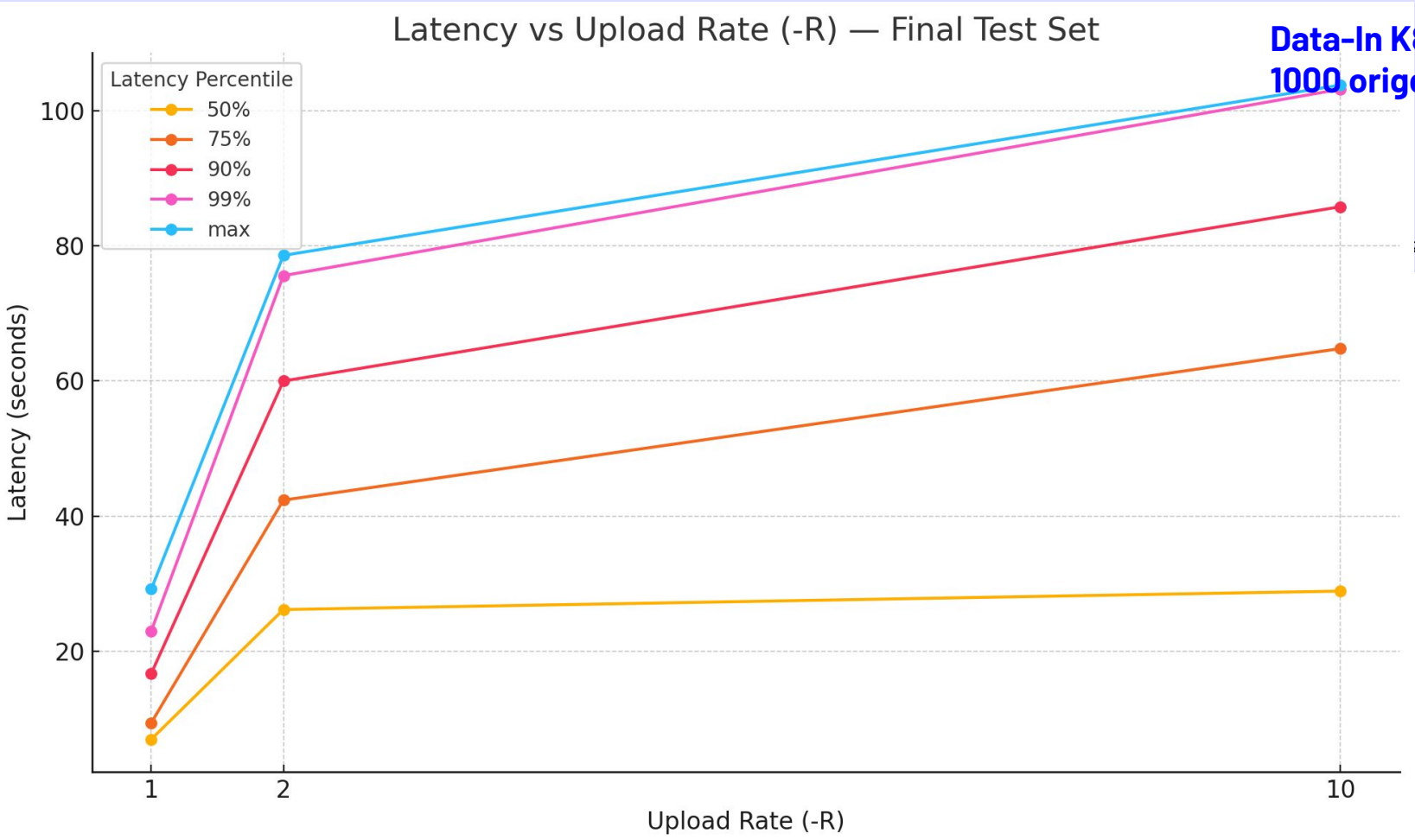
Requested by Carlo Minotti at SciCatCon 2025



100 metadata throughput increase

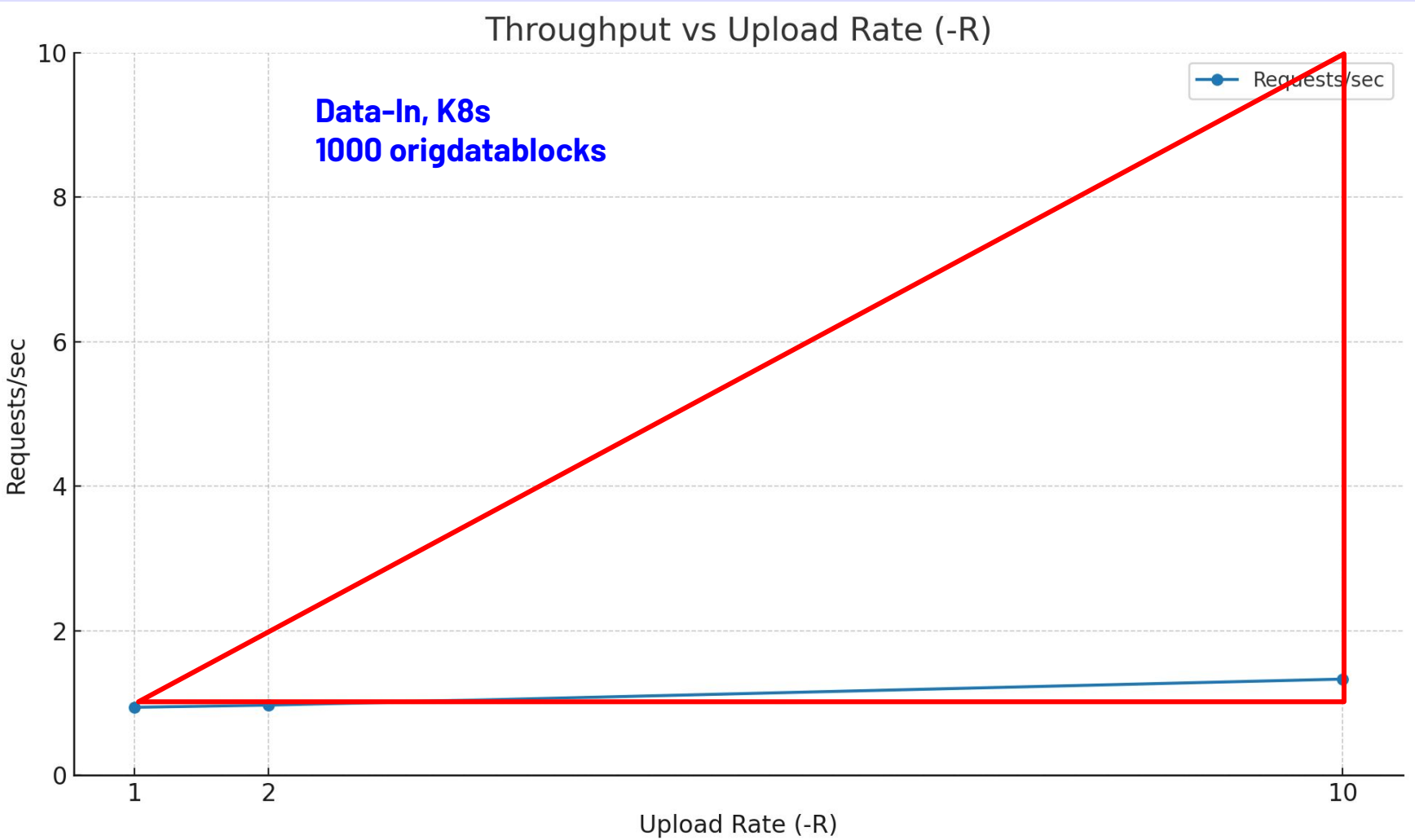
g elit.



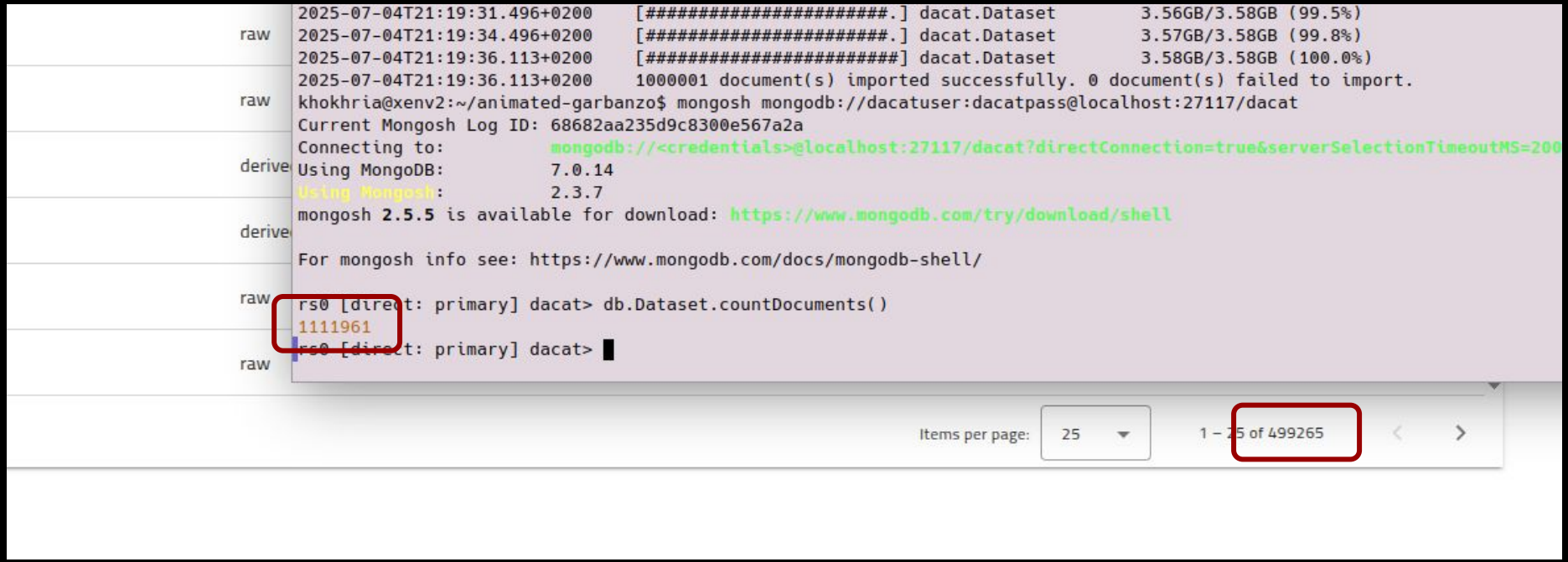


Data-In K8s
1000 origdatablocks

elit.


Throughput vs Upload Rate (-R)
Requests/sec
Data-In, K8s
1000 origdatablocks
Requests/sec
10
8
6
4
2
0
1
2
10
Upload Rate (-R)

SIDE NOTES


raw
raw
derive
derive
raw
raw
2025-07-04T21:19:31.496+0200 [########################.] dacat.Dataset 3.56GB/3.58GB (99.5%)
2025-07-04T21:19:34.496+0200 [########################.] dacat.Dataset 3.57GB/3.58GB (99.8%)
2025-07-04T21:19:36.113+0200 [########################] dacat.Dataset 3.58GB/3.58GB (100.0%)
2025-07-04T21:19:36.113+0200 1000001 document(s) imported successfully. 0 document(s) failed to import.
khokhria@xenv2:~/animated-garbanzo$ mongosh mongodb://dacatuser:dacatpass@localhost:27117/dacat
Current Mongosh Log ID: 68682aa235d9c8300e567a2a
Connecting to: mongodb://<credentials>@localhost:27117/dacat?directConnection=true&serverSelectionTimeoutMS=200
Using MongoDB: 7.0.14
Using Mongosh: 2.3.7
mongosh 2.5.5 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
rs0 [direct: primary] dacat> db.Dataset.countDocuments()
1111961
rs0 [direct: primary] dacat>
Items per page:
25
1 – 25 of 499265

Thanks

Directed by
Igor Khokhriakov aka Ingvord